Claims

1. A knowledge map representation of selected discrete perspectives of a domain of knowledge, said knowledge map useable by a knowledge instance classification computer program to classify instances of knowledge according to said discrete perspectives, said representation comprising:

a plurality of separate taxonomies, each separate taxonomy representing one of said discrete perspectives of the knowledge domain;

wherein each taxonomy is organized as a graph of nodes, connected by edges, each node of said graph corresponding to a conceptual area within the discrete perspective that the taxonomy represents and each edge of said graph representing a parent-child relationship between the conceptual areas to which the nodes connected to that edge correspond.

2. The knowledge map of claim 1, wherein said graph is a directed acyclic graph.

3. The knowledge map of claim 1, wherein said graph is a hierarchical graph.

4. The knowledge map representation of claim 1, further comprising at least one knowledge instance associated with at least one node of said taxonomies;

whereby the at least one association is a classification of the knowledge instance within the knowledge domain.

5. The knowledge map representation of claim 1, wherein each of said taxonomies is one of:

a topic taxonomy, in which each node of the topic taxonomy represents one or more topics and the association of a knowledge container with a node of said topic taxonomy indicates that the content of the knowledge container is about the topic represented by that node;

a filter taxonomy, in which each node of the filter taxonomy represents meta-data which are characteristics of knowledge containers that cannot be readily derived from the content of the knowledge container and the association of a knowledge container with a node of said filter taxonomy indicates that the knowledge container has the characteristic represented by that node; or

a lexical taxonomy, in which each node of the lexical taxonomy represents concepts in the knowledge domain that are identifiable by one or more specific words or phrases and the association of a knowledge container with a node of said lexical taxonomy indicates that the knowledge container has one or more instances of the words or phrases indicative of the concept represented by that node.

6. The knowledge map representation of claim 1, wherein each of said taxonomies is one of:

a process taxonomy, in which each node of the process taxonomy represents a step in one or more business processes and the association of a knowledge container with a node of said process taxonomy indicates that the content of the knowledge container is pertinent to the step represented by the node;

an environment taxonomy, in which each node of the environment taxonomy represents at least one entity;

a diagnostic taxonomy, in which each node of the diagnostic taxonomy represents at least one symptom of a problem and the association of a knowledge container with a node of said diagnostic taxonomy indicates that the content of the knowledge container describes a method to address that symptom;

a human characteristics taxonomy, in which each node of the human characteristics taxonomy represents attributes (e.g., address, height, weight, etc.), and the association of a knowledge container with a node of the human characteristics taxonomy indicates that the content of the knowledge container concerns the attribute represented by the node;

an entitlement taxonomy, in which each node of the entitlement taxonomy represents an access control level of permission for viewing the content or accessing the resources of knowledge containers and the association of a knowledge container with a node of said entitlement taxonomy indicates that the knowledge container is to have the access control level specified by that node; or

a standard taxonomy, in which each node of the standard taxonomy represents topics, and the association of a knowledge container with a node of the standard taxonomy indicates that the content of the knowledge container concerns the topic represented by the node.

7. A knowledge map representation as in claim 6, wherein the at least one entity is a person, place, organization, product, family of products, or a customer segment.

8. The knowledge map representation of claim 1, wherein a taxonomic distance function is associated with each pair of nodes of a taxonomy and is a function of the graphical representation of the taxonomy.

9. The knowledge map representation of claim 8, wherein each pair of nodes includes a parent node and a child node, and wherein the taxonomic distance function for the pair of nodes in the direction from the parent node to the child node is different than for the direction from the child node to the parent node.

10. The knowledge map representation of claim 8, wherein the taxonomic distance function for each pair of nodes of a taxonomy at least partially depends on how deep in the taxonomy are the pair of nodes.

11. The knowledge map representation of claim 8, wherein the taxonomic distance function includes parameters incorporated manually by a human user, whereby the taxonomic distance function accounts for human knowledge about a semantic distance between concept nodes in the taxonomy.

12. The knowledge map representation of claim 11, wherein the manually incorporated parameters are represented in an editable table that facilitates computing the taxonomic distance function for a particular pair of nodes.

13. The knowledge map representation of claim 8, wherein the taxonomic distance function for a pair of nodes of a taxonomy at least partially depends on that taxonomy

14. The knowledge map representation of claim 1, wherein the knowledge map includes a plurality of knowledge map regions, wherein each knowledge map region is a group of nodes collectively representing a coherent subdomain of knowledge.

15. The knowledge map representation of claim 14, wherein all the nodes in a particular knowledge map region are in the same taxonomy.

16. The knowledge map representation of claim 15, wherein:

taxonomic distance is a function of the graphical representation into which the taxonomy is organized;

the knowledge map region is centered about a particular central node; and

the nodes that are members of the region are those nodes having the least taxonomic distance from the particular central node.

17. The knowledge map representation of claim 15, wherein:

knowledge containers are associated with at least some of the nodes; and

the nodes that are a member of the region have a similarity of vocabulary in the content of knowledge containers associated with the nodes.

18. The knowledge map representation of claim 15, wherein:

the knowledge map region is centered about a particular central node;

knowledge containers are associated with at least some of the nodes; and

the nodes that are members of the region are those nodes having the least taxonomic distance from the particular central node and also for which there is similar vocabulary in the content of knowledge containers associated with the nodes.

19. A knowledge container, including:

an indication of an object; and

at least one tag, wherein each tag associates the object to a knowledge map representation of a discrete perspective of a domain of knowledge.

20. The knowledge container of claim 19, wherein the object is one of content and resources.

21. The knowledge container of claim 19, further including administrative meta-data, comprised of structured information about the object.

22. The knowledge container of claim 19, wherein the indication of the object is the object itself.

23. The knowledge container of claim 19, wherein the indication of the object is a pointer to the object.

24. The knowledge container of claim 22, wherein the knowledge container includes:

marked content that is a textual representation of the object;

selective demarcation of regions of the textual representation of the object; and

a plurality of indicators of the nature of the content.

25. The knowledge container of claim 19, wherein each tag includes a weight indication representing a strength of association of the knowledge container to a particular node.

26. The knowledge container of claim 21, wherein the administrative metadata contains a description of the method used to assign the knowledge container to a particular node, including:

SME designation;

autocontextualization;

source mapping based on where the knowledge container came from; and

dialog response.

27. The knowledge container of claim 19, wherein said at least one tag is associated with nodes from a single taxonomy.

28. The knowledge container of claim 19, wherein said at least one tag is associated with nodes from a plurality of taxonomies.

29. The knowledge container of claim 19, wherein the object indicates a person's interests, information needs, and entitlements.

30. The knowledge container of claim 29, wherein the tags for the knowledge container include a weight representing:

a strength of the person's interest or information need;

relevancy to a question; and

expertise of a provider.

31. The knowledge container of claim 30, wherein the tags for the knowledge container associate the knowledge container with various portions of the knowledge map.

32. The knowledge container of claim 29, wherein the person's entitlements are represented as tags to nodes of an entitlement taxonomy.

33. The knowledge container of claim 19, wherein the knowledge container is represented by a markup language such that it is displayable using template-based automated processing.

34. A method of processing at least one tag to generate a summary of a knowledge container, comprising the steps of:

generating a natural language template based on at least one tag stored inside the knowledge container; and

merging content from the knowledge container and the tagged concept-nodes into the template.

35. An autocontextualization method to automatically associate a knowledge container with a knowledge map having a plurality of taxonomies representative of selected discrete perspectives of a knowledge domain, each taxonomy having nodes corresponding to a conceptual area within the discrete perspective that the taxonomy represents, the autocontextualization method comprising:

using a feature recognizer to determine features of the knowledge container ;

employing a classification system to classify the knowledge container based on the determined features;

generating a preliminary list of nodes to which the knowledge container may be associated; and

determining a weight indicating a strength of association therewith.

36. The autocontextualization method of claim 35, further including the steps of:

truncating nodes from the preliminary list based on the strength of association indicated by the weights; and

generating an indication that the remaining nodes are associated with the knowledge container.

37. The autocontextualization method of claim 35, further including:

following the classifying step, adjusting the weights determined by the classification system by applying an inference engine based on a set of rules regarding relationships between the nodes.

38. The autocontextualization method of claim 35, further including:

following the classifying step, adjusting the preliminary list of nodes generated by the classification system by applying an inference engine based on a set of rules regarding relationships between nodes:

39. The autocontextulization method of claim 35, wherein the feature recognizer recognizes as features at least some of:

dates;

times;

numbers;

monetary amounts;

people's names;

organization names;

product names;

company names;

technical terminology;

noun phrases;

verb phrases; and

syntactic relationships.

40. The autocontextualization method of claim 35, wherein the step of generating a preliminary list of nodes further includes the step of identifying the features within the content most relied upon by the classifier in making the classification.

41. An organization of a contiguous entity of knowledge, comprising:

a plurality of knowledge containers, each knowledge container having an indication of a constituent portion of the entity of knowledge, each constituent portion of the entity relating to a different topic; and

at least one tag associated with said knowledge container, wherein the tag represents an association of a constituent portion of the knowledge container to a concept node.

42. The organization of claim 41, further comprising at least one link associated with a first knowledge container, wherein said at least one link associates said first knowledge container to at least a second knowledge container.

43. The organization of claim 41, wherein each of the knowledge containers are subordinate knowledge containers, further including:

a master knowledge container that includes an indication of the entire entity of knowledge, wherein each of the subordinate knowledge containers include a link to the master knowledge container.

44. A method of processing a query to identify a particular knowledge container, associated with a knowledge map, that is relevant to the query, wherein the knowledge map includes at least one taxonomy representing a discrete perspective of a knowledge domain, wherein the at least one taxonomy is organized into a group of nodes, the nodes representing conceptual areas within the discrete perspective, and wherein the nodes have an indication of knowledge, including the particular content associated therewith, said method comprising the steps of:

(a) processing the query to identify nodes of the taxonomies within the knowledge map that are potentially relevant conceptual areas;

(b) identifying knowledge map regions surrounding at least one of the identified nodes;

(c) performing a content-based retrieval over the knowledge containers associated with the nodes in each identified region, to retrieve an ordered list of potentially relevant knowledge containers, where each retrieved knowledge container is assigned a numerical relevance score representing a quality of association between the retrieved knowledge container and the query;

(d) combining the ordered lists for the identified regions into a single re-ordered list, based on calculating the quality of associations between the knowledge container in the list, the knowledge map, and the query; and

(e) returning as a result the re-ordered list of the retrieved knowledge containers.

45. The method of claim 44, further including the step of returning the potentially relevant nodes and knowledge map regions.

46. The method of claim 44 wherein the content based retrieval step operates upon one content-based search engine index for all knowledge containers associated with nodes of the knowledge map.

47. The method of claim 44 in which the content-based retrieval step operates on at least one distinct content-based search engine index per region, where each index indexes or points to a subset of the knowledge containers associated with nodes of the knowledge map.

48. The method of claim 47 , wherein for each concept node in at least some of the taxonomies, the knowledge containers whose content is associated with those nodes are indexed by a distinct index.

49. The method of claim 47, wherein in the subset of knowledge containers in each index have similarity of vocabulary.

50. The method of claim 49, wherein the subsets of knowledge containers in each index are formed by steps of:

aggregating the content indicated by knowledge containers associated with each node into a single block of content;

grouping the blocks together based on vocabulary occurring within the blocks, using a text clustering system; and

grouping those knowledge containers whose content comprises the forming the knowledge containers from which the blocks in a group originate into a distinct subset.

51. The method of 44, wherein the content-based retrieval step is performed over a group of indexes for each knowledge map region, wherein the group of indexes for a particular region is based on indexes for nodes in that knowledge-map.

52. The method of claim 44, wherein the query processing step further includes the step of augmenting the set of identified nodes with additional nodes as input to the query process.

53. The method of claim 47, wherein the content-based retrieval step further includes:

performing an additional search over an index for all knowledge containers associated with concept nodes in the knowledge map.

54. The method of claim 44, wherein the list combining step includes the following steps:

modifying the numeric relevance scores; and

combining the ordered lists into the single reordered list based on the modified relevance scores;

wherein the numeric relevance score for a knowledge container in a particular knowledge map region is modified at least partially based on a quality measure for that knowledge map region.

55. The method of claim 54 wherein the quality measure for a particular knowledge-map region is derived from a quality measure for each of the potentially relevant concept nodes around which the knowledge-map region surrounds.

56. The method of claim 55, wherein the quality measure for a potentially relevant concept node is based on the weight value determined in the query process step when identifying a node for a potentially relevant conceptual area.

57. The method of claim 55, wherein the quality measure for a node for a potentially relevant conceptual area is based on a weight for that node determined in the query process step.

58. The method of claim 54, wherein the numeric relevance score for a particular knowledge container is adjusted based on a quality measure for that knowledge container.

59. The method of claim 54, wherein the quality measure for a particular knowledge container is based on weights of association of the knowledge container with nodes of the taxonomies.

60. The method of claim 54, wherein the quality measure for a particular knowledge container is based at least in part by how many knowledge map regions with which the knowledge container has associated nodes.

61. The methods of claim 54, wherein the quality measure for a particular knowledge container is dependent on a taxonomic distance between the nodes in the knowledge map with which the knowledge container is associated and nodes in the knowledge map with which the query is associated.

62. The method of claim 54, wherein the query is a present query, and wherein the quality measure for a particular knowledge container is based at least in part on a previously-determined overall quality score for the knowledge container based on from users presented with the knowledge container in response to previous queries.

63. The method of claim 44, wherein the query includes taxonomic restrictions limiting the areas of the knowledge map from which a knowledge container is returned in response to the query,

64. The method of claim 63, wherein the taxonomic restrictions include:

a) a restriction that all knowledge containers returned must be associated with nodes in a particular one or more of the taxonomies;

b) a restriction that all knowledge containers returned must be associated with particular nodes;

c) a restrictions that all knowledge containers returned must be associated with nodes either at or taxonomically under a particular node or nodes; and

d) a boolean combination of the restrictions a), b) and c).

65. The method of claim 64, where said taxonomic restrictions further include a restriction that all knowledge containers returned must be tagged to concept-nodes either at or within a particular taxonomic distance of a particular concept-node or nodes.

66. The method of claim 64, where said taxonomic restrictions further include:

a) a restriction that all knowledge containers returned may not be associated with nodes in a particular one or more of the taxonomies;

b) a restriction that all knowledge containers returned may not be associated with particular nodes;

c) a restrictions that all knowledge containers returned may not be associated with nodes either at or taxonomically under a particular node or nodes; and

d) a boolean combination of the restrictions a), b) and c).

67. The method of claim 44, further including a step of processing administrative meta-data constraints to limit the knowledge containers included in the result, the administrative meta-data constraints including at least one of:

names of authors of the knowledge containers;

date ranges for creation date of the knowledge containers;

date ranges for last modified date of the knowledge containers;

date ranges for expiration date of the knowledge containers;

words or phrases which must be present in the title of the knowledge containers;

name of publication or source in which the knowledge containers originally appeared; and

name of customers for which the knowledge containers were originally prepared.

68. The method of claim 65, further including the step of constructing the taxonomic restrictions.

69. The method of claim 65, wherein said constructing step is further comprised of the step of manually interacting with a graphical display of the knowledge map to indicate desired taxonomic restrictions.

70. The method of claim 65, wherein the interfacing step includes the step of receiving a textual query from the user.

71. The method of claim 65, wherein indications of knowledge experts are associated with nodes for which the conceptual areas represented by the nodes are with the experts' area of expertise, and

wherein information about the experts may be included as part of the result of processing the query.

72. The method of claim 44, further including the following steps:

receiving input from a user as to the suitability of particular portions of the returned result;

modifying the query in response to the input; and

repeating steps (a) - (e), using the modified query.

73. The method of claim 44, further comprising the step of generating clarifying questions based on the nodes for potentially relevant knowledge containers, wherein the input is provided at least partially in response to answers from a user to the clarifying questions.

74. The retrieval method of claim 44, further comprising the step of generating suggested additional terms for the query based on the nodes for potentially relevant knowledge containers, wherein the query is modified in response to a user choosing from the additional terms.

75. The retrieval method of claim 44, further comprising the steps of:

generating parameterized questions from which a user can interactively construct a taxonomic restriction to limit the areas of the knowledge map or construct a query from which result knowledge content is returned in response to the query, said parameterized questions including:

a boolean taxonomy-restriction expression, where the concept nodes in the expression are replaced with variables;

text of a previously composed question comprised of a plurality of text selection-list boxes for each variable within the boolean taxonomy-restriction expression, wherein each selection-list box holds lists of names or descriptions of concept-nodes that are potential values for the variable;

said lists being assembled using the names or descriptions of concept-nodes returned by the retrieval mechanism in the previous step of the dialog, possibly augmented with other nearby concept-nodes from the same taxonomies;

said selection-list boxes optionally having pre-selected as the default choice for the user the specific concept-nodes returned by the retrieval mechanism in the previous step of the dialog, such that when a user selects concept-nodes for each selection-list box within the parameterized question, the boolean taxonomy-restriction expression is instantiated by replacing each of its variables with the corresponding selection-list box selection, and the resulting taxonomic restriction is added to the user's query for the subsequent step of the dialog.

76. The method of claim 44, wherein the knowledge container includes other intellectual content or an indication of a person who has knowledge contact is associated.

77. The retrieval method of claim 44, wherein:

some of the content associated with the nodes of the knowledge map include an indication of a user and the user's interests; and

at least some of the steps of the retrieval process account for the user's interests.

78. The retrieval method of claim 77, wherein the steps that account for the user's interests include the list combining step.

79. The retrieval method of claim 78, wherein the numerical relevance scores are modified based on a correlation between the user's interests and the nodes with which the retrieved knowledge container is associated.

80. The retrieval method of claim 52, wherein the method is initiated from a user application, and wherein information about the user application is provided in the form of concept nodes added to the query.

81. The retrieval method of claim 69, wherein the process is initiated from a user application, and wherein information about the user application is provided as the taxonomic restrictions.

82. The knowledge retrieval process of claim 44, wherein the process is initiated from a user application, and wherein the list combining step operates based on information about the user application.

83. The knowledge retrieval process of claim 44, wherein the list combining step operates at least in part based on an identification of nodes of the knowledge map by a user.

84. A method of identifying a knowledge container associated with a knowledge map, wherein the knowledge map includes at least one taxonomy representing a discrete perspective of a knowledge domain, wherein the at least one taxonomy is organized into a group of nodes, the nodes representing conceptual areas within the discrete perspective, and wherein the nodes have an indication of knowledge, including the particular content associated therewith, said method comprising:

processing information about a user to identify nodes in the taxonomy that represent conceptual areas previously indicated to be of interest to a user;

identifying knowledge map regions surrounding the at least one of the identified nodes; and

performing a content-based retrieval over the knowledge containers associated with the nodes in each identified region, to retrieve an ordered list of potentially relevant knowledge containers, where each retrieved knowledge container is assigned a numerical relevance score representing a quality of association between the retrieved knowledge container and the customer information.

85. The knowledge container of claim 29, wherein the indication of the user's interests and information needs includes a query for use by a retrieval method to retrieve objects mapped to the knowledge map.

86. The method of claim 84, wherein the information about the customer is processed automatically with any action by the user, and wherein at least one portion of the knowledge container of the re-ordered list is displayed to the user.

87. A method for constructing a knowledge map from a corpus of knowledge containers, said method comprising:

identifying a set of root nodes for proposed discrete taxonomies to represent facets of the domain of knowledge;

extracting terms and features from the corpus of knowledge containers;

assigning the terms amongst the proposed discrete taxonomies;

constructing each taxonomy from the terms ascribed to that taxonomy and the corpus of knowledge containers; and

testing and refining each constructed taxonomy using a text classification system.

88. The method of claim 87, wherein the constructing step includes,

clustering the set of terms ascribed to each taxonomy into multiple groups based on the usage and collocation in the text of the corpus;

organizing the term clusters into a hierarchical taxonomy based on statistical correlations among the term clusters, each term cluster becoming a concept-node in the hierarchical taxonomy;

assigning each concept-node within the generated taxonomy an appropriate name; and

manually reviewing and altering the generated taxonomy.

89. The method of claim 88, wherein the clustering step utilizes:

correlations between terms in each cluster; and

the number of terms in each cluster;

wherein a sequence of clustering steps are carried out with each subsequent step attempting to discover additional clusters and using the highest level clusters as higher level concept-nodes.

90. The method of claim 88, wherein the name assigning step includes assigning concept-node names to each term cluster based on human manual review by a human expert within the knowledge domain.

91. The method of claim 90, wherein the manual review step includes at least one of:

adding concept-nodes to the generated taxonomy;

removing concept-nodes from the generated taxonomy;

moving concept-nodes within the generated taxonomy to different locations; and

adding or removing terms from the term clusters associated with concept-nodes.

92. The method of claim 87, wherein the testing and refining step includes using a trainable text classification system, to perform:

for a subset of knowledge container in the corpus, creating a training set by manually identifying the concept-nodes from the newly generated taxonomy that correspond to topics that appear within the content of the knowledge container;

for a subset of knowledge containers in the corpus, creating a test set by manually identifying the concept-nodes from the newly generated taxonomy which correspond to topics that appear within the content of the knowledge container;

training the text classification system by using the content of the knowledge containers identified for each concept-node in the training set as example data for the concept node;

generating a test on training set report indicating how well the trained text classification system's classification of the knowledge containers in the training set for each concept-node within the taxonomy matches the manually identified classification;

based on the test on training set report, refining the taxonomy by at least one of:

adding concept-nodes to the generated taxonomy;

removing concept-nodes from the generated taxonomy;

moving concept-nodes within the generated taxonomy to different locations; and

adding or removing knowledge containers from the manually-identified training set of knowledge containers corresponding to each concept-node;

generating a test on test set report indicating how well the trained text classification system's classification of the knowledge containers in the test set for each concept-node within the taxonomy matches the manually identified classification; and

based on the test on test set report, refining the taxonomy by one or more of:

adding concept-nodes to the generated taxonomy;

removing concept-nodes from the generated taxonomy;

moving concept-nodes within the generated taxonomy to different locations; and

adding or removing knowledge containers from the manually-identified training set of knowledge containers corresponding to each concept-node.

93. A method for constructing a knowledge map from a corpus of knowledge comprising a plurality of knowledge containers, said method comprising the steps of:

(a) identifying a plurality of taxonomies that represent major subject areas found in the plurality of knowledge containers;

(b) distributing each of the plurality of knowledge containers into at least one of the plurality of taxonomies;

(c) for each taxonomy, identifying a set of concept-nodes that represent major themes found in the plurality of taxonomy knowledge containers;

(d) for each concept-node:

distributing the plurality of knowledge containers into each of the plurality of concept-nodes; and

designating the concept-node as a taxonomy for the portion of the knowledge map to be constructed;

(e) repeating steps (c) and (d) for each taxonomy until the knowledge map cannot be further expanded; and

(f) testing and refining each concept-node of said knowledge map using a text classification system.

94. The method of claim 93, wherein the step of testing and refining each concept-node includes at least one of the following steps:

adding concept-nodes to the generated taxonomy;

removing concept-nodes from the generated taxonomy; and

moving concept-nodes within the generated taxonomy to different locations.

95. The method of claim 93, wherein the testing and refining step includes using a trainable text classification system, to perform:

for a subset of knowledge containers in the corpus, creating a training set by manually identifying the concept-nodes from the newly generated taxonomy that correspond to topics appearing within the content of the knowledge container;

for a subset of knowledge containers in the corpus, creating a test set by manually identifying the concept-nodes from the newly generated taxonomy that correspond to topics appearing within the content of the knowledge container;

training the text classification system by using the content of the knowledge containers identified for each concept-node in the training set as example data for the concept node;

generating a test on training set report that indicates how well the trained text classification system's classification of the knowledge containers in the training set for each concept-node within the taxonomy matches the manually identified classification.

96. The method of claim 95, further comprising the following steps of:

refining the taxonomy by at least one of:

adding concept-nodes to the generated taxonomy;

removing concept-nodes from the generated taxonomy;

moving concept-nodes within the generated taxonomy to different locations; and

adding or removing knowledge containers from the manually-identified training set of knowledge containers corresponding to each concept-node; and

generating a test on test set report indicating how well the trained text classification system's classification of the knowledge containers in the test set for each concept-node within the taxonomy matches the manually identified classification.

97. The method of claim 96, further comprising the step of refining the taxonomy by at least one of:

adding concept-nodes to the generated taxonomy;

removing concept-nodes from the generated taxonomy;

moving concept-nodes within the generated taxonomy to different locations; and

adding or removing knowledge containers from the manually-identified training set of knowledge containers corresponding to each concept-node.